Expanding theory-based evaluation: incorporating value creation in a theory of change[1]

Julian King[2]

Abstract

Evaluations of policies and programs often use a theory of change to articulate how the intervention is intended to function and the mechanisms by which it is supposed to generate outcomes. When an evaluation includes cost and efficiency considerations, economic and other concepts can be added to a theory of change to articulate a *theory of value creation* that articulates the mechanisms by which the intervention should use resources efficiently, effectively and create sufficient value to justify the resource use.

This paper introduces some theories of value creation that are often implicit in program designs. Making these theories explicit can support clearer evaluative thinking about value for money - including specification of criteria and standards that are aligned with the theory, methods of inquiry that test the theory, and well-reasoned judgements that answer evaluative questions about value for money. Implications for evaluation practice will be discussed.

Introduction

It has been argued for many years that program evaluators should consider cost when evaluating programs and policies, including greater use of economic methods of evaluation (Gargani, 2017; Herman, Avery, Schemp, & Walsh, 2009; Julnes, 2012; Levin, 1987; Scriven, 2013; Yates, 1996). Conversely, there are principles and practices that economists conducting cost-benefit and similar analyses can usefully draw from the field of program evaluation. For example, it has been argued that economists should incorporate explicit evaluative reasoning, evaluative thinking, and program evaluation standards in their work (King, 2019b).

Economics and evaluation sit on different sides of a disciplinary divide, for the most part peaceably and without taking much notice of one another, as alternative ways to evaluate programs and policies. But they can be combined (King, 2017). Doing so provides opportunities to integrate insights from economic evaluation with broader values and evidence. For example, value for money (VFM) – in essence, good resource use – can be evaluated using a model that integrates multiple criteria (such as efficiency and equity) with mixed methods evidence. The model has been tested in case studies, providing proof of concept (King, 2019b).

The focus of this article is on the role of program theory as a tool to facilitate the integration of evaluative and economic thinking – and on the potential to augment a theory of change (how we suppose change happens) by articulating a *theory of value creation* (how we suppose value is created).

---

[1] King, J. (2021). Expanding theory-based evaluation: incorporating value creation in a theory of change. *Evaluation and Program Planning.* https://doi.org/10.1016/j.evalprogplan.2021.101963

[2] Director, Julian King & Associates Limited www.julianking.co.nz

The article proceeds by briefly introducing a real world program that will be used to illustrate the arguments proposed. It then introduces three foundational concepts: economic methods of evaluation, evaluative reasoning, and theories of change. It is argued that: i) economic methods of evaluation offer unique insights but are insufficient on their own to comprehensively determine the value of a social program; ii) the use of evaluative reasoning and mixed methods permits a broader approach that can incorporate strengths of economic evaluation while compensating for some of its limitations; and iii) a theory of change is an important foundational element for such an approach – and it could be extended by explicitly articulating the mechanisms by which the intervention should use resources efficiently, effectively and create sufficient value to justify the resource use. To address this opportunity, a new innovation is proposed – a theory of value creation. Finally, implications for evaluation practice are discussed.

Like Minds, Like Mine: a public awareness program to reduce stigma and strengthen social inclusion

The Like Minds, Like Mine program[3] aims to reduce stigma, increase social inclusion and end discrimination towards people with experience of mental distress. Funded by the New Zealand Government and delivered by *Te Hiringa Hauora*, the Health Promotion Agency, the program celebrated its 20th anniversary in 2017. In its current iteration the program comprises a mix of public awareness campaigns and community projects aimed primarily at influencing people, environments and infrastructures that have the power to exclude people with mental distress (Health Promotion Agency, 2014).

Multiple organizations are funded under the Like Minds, Like Mine umbrella to contribute to its delivery. The Mental Health Foundation provides national communications, including e-newsletters, website, social media, and media engagement. It also administers a community grants fund, creative grants and media grants, supporting initiatives that contribute to reducing stigma and discrimination and promoting social inclusion. Non-government organizations (NGOs) are contracted to deliver anti-stigma and discrimination education projects in priority settings (social housing, education, police, health care and workplaces). Another project, aimed at growing mental health leadership, supports people to use their own story of mental distress and recovery to reduce discrimination in their communities (Health Promotion Agency, 2014).

Increasingly, there is governmental scrutiny of policies and programs to determine whether they provide VFM. This is challenging in the Like Minds, Like Mine context for several reasons. Among these is the nature of the intended outcome of "a socially inclusive New Zealand that is free of stigma and discrimination towards people with mental illness", as expressed in the program's outcomes framework (Ministry of Health and Health Promotion Agency, 2014, p. 12). Achieving this outcome is best understood as a journey. It is not a final destination that will ever be reached; rather it is part of an ongoing, dynamic set of social processes, as reflected in the 20-year history of the program.

Moreover, the social systems that support inclusion (such as social and cultural networks, positive relationships, cohesive communities, and social participation) are dynamic, complex, and much larger than Like Minds, Like Mine. The ambition of the program is to make a

[3] https://www.likeminds.org.nz

meaningful contribution to ongoing social change, with relatively modest resources, through strategic interventions targeting specific parts of the system.

The Like Minds, Like Mine outcomes framework suggested that concrete signs of increased social value would include positive changes in: workplace policies, structures and cultures; media portrayals of people with mental distress; and community action to increase social inclusion. The program is also underpinned by a set of *guiding principles*, which are expressed at a broad level in the National Plan (Ministry of Health and Health Promotion Agency, 2014).[4] An evaluation of the program for the period 2017-2019 included collaborative work with stakeholders to articulate the features of good design, delivery and immediate outcomes (early signs of behavior change) as embodied in the guiding principles (McKegg, Crocket, Were & King, 2018). The principles served a dual purpose as a navigation aid for projects and organizations contributing to the program, and as a framework for evaluating the program (Patton, 2017). The evaluation of Like Minds, Like Mine spanned the program's design, delivery, outcomes and VFM. The focus of this paper is the VFM component.

When evaluating VFM, economic methods of evaluation are often used. Indeed, such methods have previously been applied to contribute to understanding the value of Like Minds, Like Mine (Vaithianathan & Pram, 2010). It is to these methods that our attention will now briefly turn, in order to both illustrate their strengths, and why they may not be sufficient on their own to evaluate VFM.

Economic methods of evaluation

Economic methods of evaluation compare the efficiency of alternative policies, programs or interventions by systematically identifying and valuing their costs and consequences (Drummond, Sculpher, Torrance, O'Brien, & Stoddard, 2005). Costs and consequences are set out as time series and adjusted using a *discount rate* that converts future value into present value. Discounting reflects the *opportunity cost* of resource use – that is, the foregone value that could have been generated by an alternative investment. The higher the discount rate used, and the further in the future a cost or consequence occurs, the lower its present value.

Different economic methods can be distinguished by the units of measurement used to value consequences. An illustration of the different methods – including cost-benefit, cost-effectiveness, and cost-utility analysis, and their corresponding indicators are summarized in an infographic by Vaca and King (2020).[5] Of particular relevance to Like Minds, Like Mine, cost-benefit analysis (CBA) values both costs and benefits in monetary units, enabling them to be directly compared. If the present value of benefits exceeds the present value of costs – in other words, if net present value (NPV) is greater than zero – it is concluded that the investment is worthwhile.

CBA can be used as a method of financial analysis, to evaluate the profitability of financial investments (Levy & Sarnat, 1994). However, the method has theoretical roots in welfare

[4] Leadership and coordination is strong and includes people with mental distress; strong shared purpose that responds to a changing environment; multi-level approaches are used; social model of disability and human rights is the foundation of the program; public contact with people with mental distress has an equalizing effect (power of contact); program activity highlights socially inclusive behaviors (Ministry of Health and Health Promotion Agency, 2014). The evaluation added the principle that the Like Minds, Like Mine program (under the stewardship of the Health Promotion Agency) acknowledges the special relationship bvetween Māori and the Crown under the Treaty of Waitangi (McKegg et al., 2018).

[5] http://www.saravaca.com/project/outcome-efficiency-indicators/

economics and is ultimately concerned with maximizing wellbeing at a societal level (Adler & Posner, 2006). Money, in CBA, is the unit of measurement used to represent the value of things to people. In principle, anything can be valued monetarily. In practice, this is often achieved by observing people's behavior in markets or empirical studies designed to elicit their *willingness to pay* for things that they value (Drummond et al., 2005).

For example, the value expressed in the Like Minds, Like Mine outcomes framework is social value. Ultimately it includes value for people with mental distress (e.g. being able to exercise their rights as citizens and participate in society) and value for society as a whole (a more socially inclusive New Zealand). The guiding principles underpinning the program are also expressions of its value in terms of good design, delivery and immediate outcomes. In principle, monetary valuations of these intangible benefits could be measured through a *contingent valuation* study. For example, this could take the form of a survey presenting a series of hypothetical trade-offs, revealing what respondents would be willing to pay for these benefits (Drummond et al., 2005).

In practice, however, these studies can be expensive to undertake. Moreover, the choice of study design and methods can affect the estimated values (Levin & McEwan, 2001) and there is disagreement in which techniques should be used in different circumstances (Drummond et al., 2005). In practice, it has been found that CBA analysts ignore benefits that are "just too hard to estimate given current techniques" (Adler & Posner, 2006, p. 78).

Like Minds, Like Mine is a case in point. A 2010 CBA (Vaithianathan & Pram, 2010) estimated the effect of an earlier iteration of the program on employment and increased use of primary health care. These benefits are relevant dimensions of social value that could readily be measured in monetary terms. However, they are peripheral to, and arguably a poor proxy for the value of reduced stigma and increased social inclusion. Nonetheless, the analysis suggested the program had a strong prospect of generating benefits (even so narrowly defined) that exceed its costs.

Economic methods of evaluation offer unique insights that would be difficult to gain in other ways, one of which is the ability to directly reconcile costs and consequences when they are measured in the same units (Adler & Posner, 2006; King, 2015; Sunstein, 2018). However, it cannot be assumed that the benefits and costs included in the scope of a CBA give us sufficient information to determine VFM. Moreover, it is not always feasible or appropriate to employ economic methods. For example, a program might create value in ways that do not readily compute in a CBA, such as redistribution of resources (CBA measures whether society is growing a 'bigger pie', whereas social value can derive from adjusting the way the existing pie is shared), scientific value (e.g. a 'failed' experiment making a vital contribution to a body of knowledge), or cultural value (e.g. contributing to the sustainability of language and traditions in new generations of an indigenous population).

Even where economic evaluation is fit for purpose, it is often insufficient to answer a VFM question because (for example) it privileges efficiency, quantitative evidence, tangible benefits (readily valued in monetary units), and aggregative synthesis (as distinct from deliberative or other approaches to synthesis). Moreover, the basis of valuing in CBA rests on a consequentialist-utilitarian perspective that "the right course of action is the one that leads to the best consequences" (Frank, 2000, p. 77) and "consequences can be ranked according to highest total utility" (Frank, 2000, p. 79). It may be desirable to take other positions on valuing, such as deontological (ethical bottom-lines) (Adler & Posner, 2006). Economic evaluation is often too narrow in scope to answer an evaluative question about how well resources are used, which might also demand consideration of wider criteria, evidence, and ways of valuing (King, 2017).

A more capacious methodology for reconciling multiple criteria and multiple streams of evidence involves "explicit evaluative reasoning, with methods tailored to context including economic methods where feasible and appropriate. Such a model can incorporate the strengths of economic evaluation without being limited to economic criteria and metrics alone" (King, 2019b, p. 3). The process of evaluative reasoning that underpins this approach is described as follows.

Evaluative reasoning

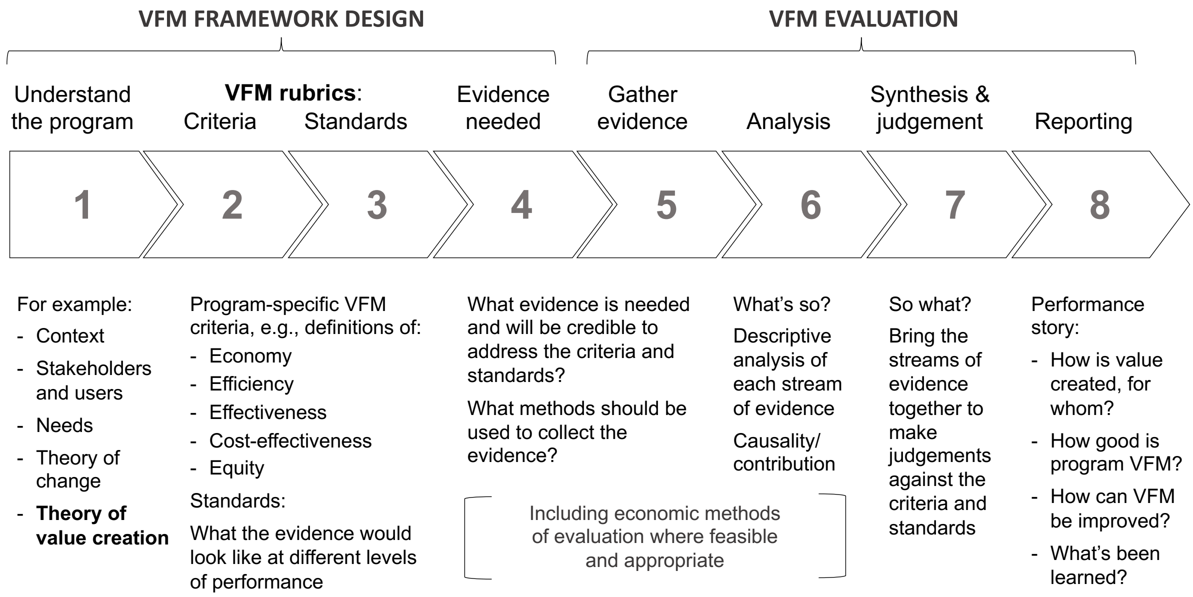
Evaluation – determining the value of things – necessitates judgements: "it does not aim simply to describe some state of affairs but to offer a considered and reasoned judgement about that state of affairs" (Schwandt, 2015). The "fundamental problem" in evaluation is how to get from evidence to evaluative conclusions (Scriven, 1995). Evaluative reasoning provides the means to make robust judgements from evidence, using a logical and traceable process (King & OPM, 2018).

Evaluative reasoning explicitly combines *values* – an expression of what matters to people (Stufflebeam, 2001; Schwandt, 2015; King, 2019b) with empirical evidence about a program, policy or intervention (Davidson, 2005; Fournier, 1995; Schwandt, 2015; Scriven, 1980; 1995). There are various different approaches to synthesizing evidence through the lens of values. Commonly, values are made explicit by specifying *criteria of merit and worth* (aspects of performance) and *standards* (levels of performance). Together, a matrix of criteria and standards is called a *rubric* (Davidson, 2005; Martens, 2018). A third dimension, importance, can also be added to a rubric, as some criteria may carry more weight than others in an evaluative judgement (King, McKegg, Oakden, & Wehipeihana, 2013).

Rubrics are just one approach to evaluative reasoning, and fall within a family of approaches that Schwandt (2015) described as "technocratic" (p. 61). Technocratic approaches to evaluative reasoning include economic methods of evaluation (King, 2019b), as well as rule-governed and algorithmic approaches (Schwandt, 2015). Examples of other models of evaluative reasoning include deliberative and tacit approaches (Ibid). These approaches are not mutually exclusive and can be combined. For example: rubrics can be used to support deliberation; tacit approaches can be used for sense-checking rubric-based judgements; and CBA can be nested within a rubric-based evaluation (King, 2019b). Irrespective of the approaches used, evaluative reasoning is integral and critical to synthesizing values and evidence, and to making an evaluative judgement based on logical argument and evidence (Davidson, 2005; Fournier, 1995; Patton, 2018; Schwandt, 2015).

The model of VFM assessment described here used rubrics as the foundation for a structured approach to evaluative reasoning (King, 2019a; King, 2019b; King & Allan, 2018; King & Guimaraes, 2016; King & OPM, 2018; Oakden & King, 2018). The practical steps involved in this process of evaluative reasoning are summarized in Figure 1. The initial steps involve developing a sound understanding of the program, then developing rubrics aligned with the program. Subsequently, rubrics provide a point of reference to determine what evidence is needed, what methods of inquiry should be used to gather evidence, and how the evidence should be interpreted to reach evaluative conclusions.

Figure 1: A stepped approach to evaluative reasoning (King, 2019b)



The contribution of a theory of change to rubric development is described as follows.

Theory of change

Any evaluation should commence by developing a sound understanding of the program including its context, stakeholders, and information needs (Patton, 2008; Scriven, 2013). These steps help to ensure the evaluation design, including the identification of criteria of merit and worth, are fit-for-purpose and aligned with the program and context. How the program is viewed "leads to a position on what evaluation approach should be taken" (Fournier, 1995, p. 23).

In these early stages of an evaluation, a theory of change can be used as a tool for facilitating and documenting a shared understanding of the program, including "what needs the program is intended to meet and how it is intended to function" (King, 2019b, p. 137). A theory of change "explains how activities are understood to produce a series of results that contribute to achieving the final intended impacts" (Rogers, 2014, p. 1). A theory of change can be developed in a participatory manner (Funnell & Rogers, 2011). This process of development can be used to build an evaluation-focused relationship among relevant stakeholders, reach an agreed understanding of the program or policy, and to foster evaluation ownership and use. A theory of change may also be used as a tool when assessing causality or contribution (Funnell & Rogers, 2011).

A theory of change can also be used to help identify a complete and coherent set of evaluation criteria that are aligned with the intended functioning of the program or policy (Davidson, 2005). For example, criteria for a process evaluation should reflect key features of activities and outputs (and relationships between them). Similarly, a theory of change offers a reference point for selecting criteria for an outcome evaluation, reflecting relevant outcomes (and their relationships to the activities and outputs of the program).

In the case of Like Minds, Like Mine, rubrics were designed to elaborate on each of the program's guiding principles. For each principle, rubrics defined what good design, delivery and immediate outcomes would look like. For example, one of the guiding principles was that people with lived experience of mental health issues from priority groups – people with severe or extreme states of mental distress, Māori, Pacific, and young people – should lead the program and its workstreams. Rubrics elaborated on this principle by specifying how it should be embodied in the program's design (e.g. lived experience leadership in decision-making and resource allocation decisions), delivery (e.g. people with lived experience having key roles), and immediate outcomes (e.g. people with lived experience being recognized as contributors and leaders in their communities and the settings they engage with) (McKegg et al., 2018).

The design of the rubrics was also informed by: i) an outcomes framework that broadly defined the intended outputs and outcomes of the program; ii) further stratification of outcomes according to levels of change (knowledge, attitudes and behavior); and iii) the specific context and content of each workstream such as target groups, settings and objectives. Together, the principles, outcomes framework, levels of change, and workstreams served as touchpoints to cohere the rubrics around the intended functioning and outcomes of the program. Collectively they sufficed to outline the program's theory of change.

To address the VFM component of the evaluation, a rubric was developed that defined features of good resource use in the program. The rubric included eight criteria, corresponding to different levels of the theory of change. To summarize, these criteria posited that the program would represent good use of resources if it creates more value than it consumes, by achieving:
1) enough of the right sorts of outcomes (outcome efficiency);
2) with the right people (social justice & equity); and that this would be enabled by
3) investing in interventions that are aligned with needs and guiding principles (relevance);
4) being a good steward of public resources (input efficiency);
5) allocating resources to deliver the right mix of outputs (allocative efficiency);
6) delivering to target groups, on time, within budget, and in alignment with guiding principles (technical efficiency);
7) working adaptively to respond to a changing environment, new opportunities and learning (dynamic efficiency); and
8) contributing to changes in awareness, knowledge, beliefs and behavior as expressed in the framework of guiding principles (effectiveness).

These eight criteria are an adaptation of the five VFM criteria commonly used in development programs: economy, efficiency, effectiveness, cost-effectiveness and equity (DFID, 2011; King & OPM, 2018).

Although the eight VFM criteria were aligned with the theory of change, they were not explicitly specified within it. In hindsight, it could have aided clarity to do so. While a theory of change articulates mechanisms and assumptions underpinning how a program might generate outcomes, it doesn't explicitly state how we suppose a program might create value. For example, it doesn't specify the manner in which inputs should be used to productively deliver outputs, and it doesn't specify the mechanisms by which the program is intended to leverage enough value to make the resource use worthwhile.

Evaluators often invest effort into understanding causality or contribution, and considerable debate within the field is concerned with different ways of doing this. However, it is less common for evaluators to think about cost, and the relationship between costs and consequences. Evaluators might pay a lot of attention to whether, and how, a set of activities generates a set of outcomes, but not to whether the outcomes are valuable enough to justify the inputs consumed, nor how the relative value of outcomes to inputs compares with alternatives. A theory of change could usefully be augmented to include a theory of value creation. Doing so could support better integration of evaluative and economic thinking in evaluation of VFM.

Theory of value creation

A theory of value creation is an extra layer in a theory of change. Just as a theory of change can be used as a tool for thinking clearly about how a program might contribute to change, and whether this is occurring, so a theory of value creation could be used as a tool for thinking clearly about how a program might create more value than it consumes, and whether it is doing so to the extent and in the ways anticipated. A theory of value creation is not separate from a theory of change, rather it is an extension of the idea. Where a change mechanism seeks to explain how a change in A causes a change in B, a value creation mechanism posits how a change in A causes a disproportionate change in B, such that value gained from the change exceeds value spent.

As a starting premise, it is proposed that any worthwhile program should create value. This value could take many different forms (e.g. cultural, social, spiritual, environmental, economic). It may involve an increase in total value for society, an increase for a particular group, or it could involve redistributing value in a way that involves making some people better off and others worse off (King, 2019b). Whatever sort of value is created, a worthwhile program should bring about not just change, but change that is valuable to somebody in some way.

Further to this premise, CBA is predicated on the notion that any program should create more value than it consumes. This may be a valid proposition in many instances, but we need to proceed cautiously. If I invest my own money for my own profit, the equation may be straightforward, but when we are investing in social change, those investing and those who are supposed to benefit are not the same people and we shouldn't assume their values are aligned (Gargani, 2017). An evaluator is typically seeking to determine the value of something to a third party, where that something is delivered by a fourth party, using resources provided by a fifth party. We need to address the question of 'value from whom, created by whom, for whom', with an understanding of the issues of power and equity underlying the investment, or our investment might fail or even do harm. We need to take great care when defining what is valuable to whom.

As a second premise, we can each intuit from our own experience that programs create value in various different ways. Even within one program, there may be multiple mechanisms of value creation. As a doorway into thinking about mechanisms of value creation, consider the concept of *leverage*. In physics, a lever is a mechanical device that amplifies the force of an input, to provide a greater output force. Using the principle of leverage, I can move a boulder that would be much too heavy for me to lift. In finance, the term 'leverage' is used analogously when borrowed money is invested to create a greater profit (analogous to the output force) than the interest on the loan (input force). To illustrate, if I borrow $10,000 from you, invest it in the sharemarket, turning it into $12,000, and return your $10,000 plus $300 interest, I will have leveraged a $1,700 profit without using any of my own money.

In a social program, the idea of leverage may be a useful concept when thinking about mechanisms the program could use to multiply value. Consider the metaphor of a ship's rudder. Small inputs, applied in just the right place, can steer a large ship. Similarly, small investments, well targeted, help ensure large investments already underway create value. Like Minds, Like Mine has finite resources, and needs to allocate them where they can make the greatest contribution to positive social change. It comprises a set of strategic interventions, aligned with leverage points identified in literature.

For example, the program targets *excluders* (people, environments and infrastructures that have the power to exclude people with mental distress) with the aim of creating conditions that support social inclusion for people with mental distress. It has a multi-level design and is targeted toward specific *benefit groups* (those who should benefit from the program, including people with severe or extreme states of mental distress, Māori, Pacific, and young people) and *recipient groups* (priority environments where social inclusion could be improved – such as police, health care, social housing, secondary and tertiary education). Guiding principles help decision makers in the program to navigate and steer projects in the intended direction.

Other value creation mechanisms can similarly be identified. Ultimately, the overarching mechanism by which Like Minds, Like Mine should (if successful) create social value, is by growing social capital. A capital asset is something that lasts across multiple accounting periods. Like the fabled goose that laid the golden eggs, the asset (the goose) remains intact and generates some form of ongoing value (the eggs). For example, an apartment is a capital asset that can earn a rental income. Financial capital can be invested to earn dividends. Like Minds, Like Mine aims to grow social capital by contributing to a socially inclusive society, free of stigma and discrimination toward people with mental distress. The payback occurs when people with mental distress are able to better exercise their rights and participate in society.

Leverage and growing social capital are not the only two value creation mechanisms seen in Like Minds, Like Mine. However, they serve to illustrate how such mechanisms could be articulated for the program. Further examples of value creation mechanisms are summarized in Table 1, together with more generalized examples from a range of settings. To help make some abstract concepts more concrete, they are described using metaphors borrowed from science, finance, economics and folklore.

Table 1: Examples of value creation mechanisms

Metaphor	Examples
Golden goose: growing a capital asset that remains intact for a long duration and generates an ongoing payback. In finance, a capital asset earns income from interest, rents or dividends.	Education is an investment in knowledge and capability that generates a return for individuals (e.g. income and quality of life) and society (e.g. innovation, productivity and taxes). Relationships, self-determination, and population health are all forms of social capital that provide ongoing paybacks in the form of wellbeing and prosperity.
Rudder: small inputs, in just the right place, steer a large ship already sailing. Similarly, small investments, well targeted, help ensure large investments already underway create value.	Investing in leadership and governance capability can be a relatively small investment to enhance the effectiveness of an organization. Prioritizing interventions, based on considerations such as relevance, probability of success, and potential size and equity of impact can increase the potential of the interventions to provide VFM.
Catalyst: an agent that causes a system to reorganize (as a hardening agent causes an epoxy resin to change from liquid to solid, or a spark changes a can of gasoline). A catalyst for social change is a small intervention that disrupts the status quo in ways that ignite change.	Market systems development programs seek to catalyse changes in the underlying structure and function of markets, so that more of the benefits from economic development and growth flow through to poor people, alleviating poverty. *Cultural fit* – congruence between the cultural contexts and values of those delivering projects and those intended to benefit from them (Goodwin, Sauni and Were, 2015) strengthens the likelihood of projects bringing about sustainable change. Shifting the locus of power so that more projects are delivered not *for* or *with* intended benefit groups, but directly *by* those groups can catalyse better outcomes (Wehipeihana, 2019).
Magnet: attracting and cohering resources around a particular issue or problem.	Challenge funds attract competitive bids for donor finance, cohering innovative minds to design solutions to a defined problem. Australia's cooperative research centres (CRCs) bring multiple universities and private organizations together to conduct scientific research with a view to accelerating commercialization and adoption of new technologies.[6]
Slingshot: harnessing energy from another object (as a rock can become a projectile by using the energy stored in an elastic band, or a spacecraft can use the gravitational field of a planet to accelerate to high speeds). A social program can harness capacities and	Selling a house through a real estate agent can facilitate ready access to a wider market of interested buyers than could be reached through self-marketing. A social program can become more effective by fostering links with other governmental and non-governmental organizations with relevant

[6] https://www.industry.gov.au/funding-and-incentives/cooperative-research-centres

capabilities that sit outside the program that increase its reach and effectiveness.	expertise and networks, who can facilitate access to target groups and help reinforce key messages.
Carrot: motivating behaviors through incentives – rewards contingent on a desired behavior.	Research and development (R&D) tax breaks encourage businesses to innovate by reducing the cost of R&D. They can act as a value creation mechanism if they motivate innovation that exceeds the value of the tax breaks. UK Aid £1 for £1 disaster response fund encourages members of the public to donate by promising to double their contribution.
Nudge: influencing behaviors by making small changes to the environment.	Opt-out systems for pension schemes and organ donation systems, mean that people are automatically enrolled, so it takes greater effort to opt out than to stay in. Text message reminders on the day of a hospital appointment increase attendance rates (Behavioural Insights Team, 2015).
Seed: providing basic ingredients to get started, and nurturing a process of growth. A seed isn't the whole plant, but contains the basic genetic material and enough food to get it started. Not all seeds will germinate, but if planted in the right soil, and nurtured with enough water and the right fertilizer, many will take hold.	In finance, seed capital is usually a relatively small sum of money to help get a new venture on its feet, so it can start putting down roots and attracting more investment. Often the seed investor takes an ongoing interest in the venture (e.g. a financial stake and a seat on the board) so they can make sure it's getting enough water and fertilizer to keep growing. Government agencies similarly provide seed funding. For example, the New Zealand government's PreSeed Accelerator Fund supports early stage technology commercialization activities with commercial benefits for the country (Ministry of Business, Innovation & Employment, 2019). School sports teams are organized for many reasons, one of which is to seed future elite players.
Safety barrier / a stitch in time: low-cost interventions to prevent expensive losses (or risks thereof); addressing a small problem early and at low cost, to avoid a larger and more expensive problem later on.	Vaccination is a low-cost intervention that prevents morbidity and mortality from infectious diseases. Seatbelts are a low-cost device that can save a life in a car crash. Screening programs can detect early stage disease so that it can be treated earlier (more effectively and often at lower cost) before it becomes more serious.
Ripple effect: intentionally creating knock-on effects, like throwing a stone to create ripples in a pond.	Train-the-trainer models are a deliberate effort to share knowledge by building an 'army' of instructors who have been taught how to deliver training to others in the intended ways.

	Viral marketing campaigns aim to spread messages about goods, services, brands or social change by encouraging people to share them through word of mouth or social media. Change management strategies often involve winning over influential people who can act as 'cheerleaders', 'champions' or 'change agents', helping overcome resistance to change in others.
Robyn Hood: creating value through social justice and equity. While the other value creation mechanisms involve leveraging additional value ('growing the pie'), this mechanism involves creating value by changing the way the existing pie is shared.	Insurance markets create value through the pooling and sharing of risk. Welfare benefits, financial aid, social security and social housing, all involve redistribution of resources to meet equity objectives. In CBA, taking a societal perspective, these 'transfer payments' have the same total value before and after redistribution so do not affect net value aside from incurring administrative costs (Drummond et al., 2005). However, such redistribution is the sole purpose and value creation mechanism in some social programs.

The value creation mechanisms listed in Table 1 are not mutually exclusive. Programs may use several mechanisms in combination, as the case of Like Minds, Like Mine illustrates. Nor is this a complete inventory of relevant mechanisms. For example, mechanisms of creating cultural value, and human rights-based mechanisms of value creation, have not been explicitly canvassed (Kate McKegg, personal communication). Nevertheless, the examples serve to illustrate some of the ways policies and programs create value, and the utility to evaluators and stakeholders of understanding such mechanisms.

Lessons learned

Extending theories of change to include theories of value creation – seeking to describe the mechanisms by which new, transformed or superior value is created from the resources consumed by a program – is significant for the field of evaluation. This extra theoretical layer represents an advance over conceptualizing programs as vehicles for 'making a difference' to programs as transformational processes that convert resources (funding, expertise, relationships, etc) into significant social value. This extra layer expands our focus from change to value, prompting us to build extra theories into evaluation including economic theories and other theories of value creation that help us get to the heart of what it means for a particular program, in a particular context, to create value.

A theory of value creation, like a theory of change, can serve a number of purposes in an evaluation. As with every aspect of an evaluation, who co-creates the understanding of value and for whom is of critical importance, especially where issues of social or cultural value are at the heart of the program. Developed in a participatory way, a theory of value creation can provide a focal point for a sense-making process with stakeholders, rights-holders and end-users to reach a clear and shared understanding of how a program functions or should function. Just as a theory of change supports clear thinking and dialogue about mechanisms of change, a theory of value creation can help evaluators and stakeholders to unpack and articulate the mechanisms that a program can use to create value. A theory of value creation can be a communication tool, articulating theory for an audience to engage with.

A theory of value creation can be an alignment tool, ensuring that the evaluation coheres with the program and its context. By explicitly specifying mechanisms of value creation, relevant and well-defined criteria of VFM can be developed. These criteria will in turn influence what evidence is collected and analysed. Case studies of VFM assessments of international development programs found that specifying criteria and standards describing 'what good VFM looks like' led to the realization that a mix of quantitative and qualitative evidence would be needed to support sound judgements (King, 2019a, 2019b; King & Allan, 2018). Similarly, mixed methods evidence was necessary to evaluate VFM of Like Minds, Like Mine (McKegg et al., 2018).

The value creation mechanisms in Like Minds, Like Mine were unpacked during the final stages of the evaluation. They added a layer of understanding that was not gained from the theory of change, nor from the eight criteria of VFM. For example, the identification of the rudder mechanism led to the insight that good VFM involves not only choosing projects aligned with needs and program objectives, but also with regard to their leverage potential – collecting evidence on 'where to press' to have the greatest impact. A key insight was that the guiding principles are part of this rudder – indeed, Michael Quinn Patton (2017) used the same metaphor when explaining why principles matter for program development and evaluation. When the intent of the program is clearly expressed in guiding principles, interventions can be designed that take the program in the desired direction. The organizations delivering the interventions can navigate by the principles when developing and adapting their delivery to be effective in different contexts. Evaluators can use the principles to determine whether the program 'ship' is sailing in the desired direction. This insight can inform decisions about the next iteration of the program. Where value creation mechanisms are explicitly identified from the outset of a program, there is a chance for enhanced understanding and more purposeful program design, delivery and evaluation.

Making value creation mechanisms explicit can support evaluative thinking, including appropriate use of economic methods of evaluation. Economic methods should be used not as gold standard approaches to evaluation but as "conditionally valid and sometimes appropriate contributors to evaluation" (King, 2017, p. 111). Having a clear theory or construct of value creation can lead to the realization that in some cases, economic evaluation is insufficient to fully answer a VFM question. For example, CBA examines the value of costs and consequences without regard to the intervening processes or mechanisms involved. CBA gives us no information about how well Like Minds, Like Mine is meeting the potential of its value proposition through the use of value creation mechanisms identified in Table 1. CBA is a tool for evaluating efficiency; it is agnostic about equity and cannot be used to determine whether Like Minds, Like Mine is allocating resources to projects that promote social inclusion in an equitable manner. However, CBA did contribute important evidence that the value of benefits narrowly defined and measured in terms of employment outcomes and utilization of primary health care are likely to exceed program costs.

It is often the case in program evaluation that long-term impacts or outcomes are challenging to measure and attribute. Indeed, evaluations are often conducted before enough time has elapsed for these outcomes even to appear. In such cases a theory of change can be used to draw vital logical links between the long-term changes the program is intended to bring about, and more immediate or short-term changes that might indicate positive travel in the intended direction. Analogously, a theory of value creation can support evaluation of processes and interim value that are assumed, with solid rationale, to contribute to long-term value. For example, in some contexts it may not be feasible to conduct any sort of economic evaluation. In such cases a theory of value creation may provide a toehold for making valid judgements from available evidence. It might not be possible to determine whether value created exceeds value consumed, but a judgement can still be made about how well a program is working toward meeting its value proposition (for an example see Spee & Oakden, 2019).

The interface between economics and evaluation can influence both sets of methods. For example, making causal and value creation mechanisms more explicit can aid clarity in an economic evaluation by providing conceptual models that aid in specifying robust economic models. To this end, the notion of value creation mechanisms may step into Realist Evaluation territory, where early work is under way exploring the potential to combine generative models of causality with economic methods of valuing (Anderson & Hardwick, 2016; King & Walsh, 2017) to address questions such as *what increases value for whom, in what circumstances*?

Conversely, economic thinking can influence evaluative thinking by lending tools and concepts that aid in evaluation design and method selection. It would be hard, for example, to comprehensively assess VFM without applying concepts like technical efficiency, allocative efficiency, and opportunity cost. Other economic concepts, such as leverage, incentives, and nudge theory, also have application in some cases.

If we can explicitly define a program's value proposition, we are better placed to evaluate it. A theory of value creation offers a new tool for thinking clearly through the design and conduct of evaluations that combine economics and evaluation to answer VFM questions.

Acknowledgements

Many thanks to Kate McKegg, Louise Were, Brian Yates and Nadini Persaud who reviewed drafts of this paper and provided helpful feedback. Value creation mechanisms outlined in Table 1 emerged from evaluation projects in collaboration with Judy Oakden, Nan Wehipeihana, Kate McKegg, Louise Were, Bonnie Scarth, Kellie Spee, Debbie Goodwin and Alicia Crocket. Any errors or omissions are the author's responsibility.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The author was a member of a team of consultants who received funding to evaluate the program that was used as a case example in this paper.

References

Adler, M.D., & Posner, E.A. (2006). *New Foundations of Cost-Benefit Analysis.* Harvard University Press.

Anderson, R., Hardwick, R. (2016). Realism and resources: Towards more explanatory economic evaluation. *Evaluation, 22*(3), 323-341.

Behavioural Insights Team (2015, October 22). *Reducing missed appointments*. Retrieved from: https://www.bi.team/blogs/reducing-missed-appointments/

Davidson, E.J. (2005). *Evaluation Methodology Basics: The nuts and bolts of sound evaluation*. Sage.

Department for International Development. (2011). *DFID's Approach to Value for Money (VfM).* DFID, UK Government.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddard, G. L. (2005). *Methods for the economic evaluation of health care programs*. Oxford University Press.

Fournier, D. M. (1995). Establishing evaluative conclusions: A distinction between general and working logic. In D. M. Fournier (Ed.), *Reasoning in Evaluation: Inferential Links and Leaps. New Directions for Evaluation, (58),* 15-32.

Frank, R. (2000). Why is cost-benefit analysis so controversial? In M.D. Adler & E.A. Posner (Eds.), *Cost-Benefit Analysis: Legal, Economic, and Philosophical Perspectives* (pp. 77-94). Chicago, IL: University of Chicago Press.

Funnell, S.C., & Rogers, P.J. (2011). *Purposeful Program Theory: effective use of theories of change and logic models*. Jossey-Bass.

Gargani, J. (2017). The leap from ROI to SROI: farther than expected? *Evaluation and Program Planning, 64,* 116-126.

Goodwin, D., Sauni, P., Were, L. (2015). Cultural fit: an important criterion for effective interventions and evaluation work. *Evaluation Matters—He Take Tō Te Aromatawai*, *1,* 25-46.

Health Promotion Agency (HPA) (2014). *Like Minds, Like Mine Work Programme.* Infographic. Health Promotion Agency.

Herman, P.M., Avery, D.J., Schemp, C.S., Walsh, M.E. (2009). Are cost-inclusive evaluations worth the effort? *Evaluation and Program Planning, 32*(1), 55-61

Julnes, G. (2012). Promoting valuation in the public interest. In G. Julnes (Ed). *Promoting Valuation in the Public Interest: Informing Policies for Judging Value in Evaluation. New Directions for Evaluation, 133,* pp 109-129.

King, J. (2015). Use of cost-benefit analysis in evaluation. Letter to the editor. *Evaluation Journal of Australasia, 15*(3), 37-41.

King, J. (2017). Using Economic Methods Evaluatively. *American Journal of Evaluation, 38*(1), 101–113.

King, J. (2019a). Combining multiple approaches to valuing in the MUVA female economic empowerment program. *Evaluation Journal of Australasia, 19*(4), 217-225.

King, J. (2019b). *Evaluation and Value for Money: Development of an approach using explicit evaluative reasoning*. (Doctoral dissertation). University of Melbourne.

King, J. (2019c). *Value for Money Criteria*. Retrieved from: https://www.julianking.co.nz/vfi/criteria/

King, J., Allan, S. (2018). Applying Evaluative Thinking to Value for Money: The Pakistan Sub-National Governance Programme. *Evaluation Matters—He Take Tō Te Aromatawai, 4*, pp. 207-235.

King, J., Guimaraes, L. (2016). Evaluating Value for Money in International Development: The Ligada Female Economic Empowerment Program. *eVALUation Matters*, Third Quarter, 2016, pp. 58-69. Africa Development Bank.

King, J., McKegg, K., Oakden, J., Wehipeihana, N. (2013). Rubrics: A method for surfacing values and improving the credibility of evaluation. *Journal of MultiDisciplinary Evaluation, 9*(21), 11-20.

King, J. & OPM (2018). *OPM's Approach to Assessing Value for Money: A Guide.* Oxford Policy Management Limited. Retrieved from: https://www.julianking.co.nz/wp-content/uploads/2018/02/OPM-approach-to-assessing-value-for-money.pdf

King, J., Walsh, B. (2017, October). *Realist Economic Evaluation.* Paper presented at the 2017 International Realist Conference, Brisbane, Australia.

Levin, H. (1987). Cost-Benefit and Cost-Effectiveness Analyses. In D.S. Cordray, H.S. Bloom, and R.J. Light (Eds). *Evaluation Practice in Review. New Directions for Program Evaluation, 34*.

Levin, H.M., McEwan, P.J. (2001). *Cost-Effectiveness Analysis. 2nd Ed*. Sage.

Levy, H., & Sarnat, M. (1994). *Capital Investment & Financial Decisions (5th Ed).* Prentice Hall.

Martens, K.S.R. (2018). Rubrics in program evaluation. *Evaluation Journal of Australasia, 18*(1), 21-44.

McKegg, K., Crocket, A., Were, L., King, J. (2018). *Like Minds, Like Mine Evaluation Framework.* Kinnect Group.

Ministry of Business, Innovation & Employment (2019). *PreSeed Accelerator Fund.* Retrieved from: https://www.mbie.govt.nz/science-and-technology/science-and-innovation/funding-information-and-opportunities/investment-funds/preseed-accelerator-fund/

Ministry of Health and Health Promotion Agency (2014). *Like Minds, Like Mine National Plan 2014-2019: Programme to Increase Social Inclusion and Reduce Stigma and Discrimination for People with Experience of Mental Illness.* New Zealand Ministry of Health.

Oakden, J. & King, J. (2018). Evaluation, in M. Tolich & C. Davidson (Eds.). *Social science research in New Zealand: An introduction* (3rd ed; pp. 180-193). Auckland University Press.

Patton, M.Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage.

Patton, M.Q. (2017). *Principles-Focused Evaluation: The GUIDE.* Guilford.

Patton, M.Q. (2018). A historical perspective on the evolution of evaluative thinking. In A.T. Vo & T. Archibald (Eds.). *Evaluative Thinking. New Directions for Evaluation, 158,* 11-28.

Rogers, P. (2014). *Theory of Change. Methodological Briefs: Impact Evaluation 2.* UNICEF Office of Research.

Schwandt, T. (2015). *Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice*. Stanford University Press.

Scriven, M. (1980). *The logic of evaluation*. Sage.

Scriven, M. (1995). The Logic of Evaluation and Evaluation Practice. *New Directions For Evaluation, 68*, 49-70.

Scriven, M. (2013, March 22). *Key evaluation checklist (KEC).* Retrieved from: http://www.michaelscriven.info/images/KEC_3.22.2013.pdf

Spee, K., & Oakden, J. (2019). *Envirolink Evaluation Report 2019*. Pragmatica Limited. Retrieved from: https://www.mbie.govt.nz/assets/envirolink-evaluation-report-2019.pdf

Stufflebeam, D.L. (2001, March). *Evaluation Values and Criteria Checklist.* Retrieved from: www.wmich.edu/evalctr/checklists

Sunstein, C. R. (2018). *The Cost-Benefit Revolution.* MIT Press.

Vaca, S., King, J. (2020, April). *Outcome efficiency indicators.* Retrieved from: http://www.saravaca.com/project/outcome-efficiency-indicators/

Vaithianathan, R., Pram, K. (2010). *Cost-Benefit Analysis of the New Zealand Mental Health Destigmatisation Programme "Like-Minds Programme".* Prepared for Phoenix Research & Ministry of Health. University of Auckland.

Wehipeihana, N. (2019). Increasing cultural competence in support of indigenous-led evaluation: a necessary step toward indigenous-led evaluation. *Canadian Journal of Program Evaluation,* 34(2), 368-384.

Yates, B.T., (2009). Cost-inclusive evaluation: A banquet of approaches for including costs, benefits, and cost-effectiveness and cost-benefit analyses in your next evaluation. *Evaluation and Program Planning, 32*(1), 52-54.